



Translation of H09-258778

[0015]

The speed corresponding to the last part of the speed waveform region L and the speed corresponding to the first part of the speed waveform region R are then compared to each other. If there is a difference between them, a speed increase/decrease process is performed, in which those speeds are gradually increased or decreased in order to obtain smooth connection of the last part of the first speech file data with the first part of the second speech file data, in other words, to obtain smooth connection of the two parts to be connected.

[0016]

Increasing the speed in the speed increase/decrease process requires entering a value smaller than 1 into each multiplying power boxes 305 and 306, or clicking an upward arrow provided besides each of the multiplying power boxes 305 and 306 using a mouse through an input device 101. Decreasing the speed in the speed increase/decrease process requires entering a value larger than 1 into each of the multiplying power boxes 305 and 306 or clicking a downward arrow provided besides each of the boxes 305 and 306 using the mouse through the input device 101. These operations result in the change of the speed waveform so as to adjust the speeds corresponding to the last part of the speed waveform region L as well as the first part of the speed waveform region R to become equal to each other. If a preview button 315 is clicked for the purpose of confirmation of the process, the information on the multiplying power of speed increase/decrease which has been set is delivered from a displayed information control section 105 to a speech compression/expansion processing section 106. This section allows the two speech-file data to be combined into one speech-file data. This combined speech file data is then played

back by the speech player 109 as required.

[0017]

After completion of the process described above, the two whole speech files are displayed respectively on the speed waveform region L and the speed waveform region R side by side. If the displayed waveforms include portions indicating significantly high and low speeds, those speeds will be subject to a uniform speed process as described later.

[0018]

The uniform speed process is to change speeds within a designated range at the same rate. The designated range can be designated by entering a value into a start time box 309 or 310 and the end time box 311 or 312 for the uniform speed process range or by dragging the mouse in the appropriate portion of the speed waveform region through the input device 101. Increasing the speed requires entering a value smaller than 1 into a multiplying power box 313 or 314 for the uniform speed process or clicking an upward arrow provided besides the box 313 or 314 for the uniform speed process by the mouse through the input device 101. Decreasing the speed requires entering a value larger than 1 into the box 313 or 314 for the uniform speed process or clicking a downward arrow provided besides the box 313 or 314 for the constant speed process by the mouse through the input device 101. These operations result in the change of the speed waveforms, and also the change of the speeds within the designated range to obtain a balance of the speed as a whole.

(8)

特

【図3】

Speed waveform region L 速度波形領域L		Speed waveform region R 速度波形領域R	
File name ファイル名	5303	File name ファイル名	5304
加速處理 speed increase/decrease process 5305		加速處理 speed increase/decrease process 5306	
倍率 倍率	<input type="text"/>	倍率 倍率	<input type="text"/>
Multiplying power Multiplying power	5307	Multiplying power Multiplying power	5308
時間 時間	<input type="text"/>	時間 時間	<input type="text"/>
Time period 等速処理	309	Time period 等速処理	310
範囲 開始 Uniform speed process S	<input type="text"/>	範囲 開始 Uniform speed process S	<input type="text"/>
Range start S 313	end <input type="text"/>	Range start S 314	end <input type="text"/>
倍率 倍率	<input type="text"/>	倍率 倍率	<input type="text"/>
multiplying power 315 プレビュー		Multiplying power 316	



JP9258778

Biblio

Page 1

Drawing



VOICE PROCESSOR

1

Patent Number: JP9258778

Publication date: 1997-10-03

Inventor(s): SHIGEOKA TOMONORI;; MATSUI ISAO

Applicant(s): HITACHI INF SYST LTD

Requested Patent: JP9258778

Application Number: JP19960069560 19960326

Priority Number(s):

IPC Classification: G10L3/02

EC Classification:

Equivalents:

Abstract

PROBLEM TO BE SOLVED: To generate one voice connected smoothly in a well balanced state from a plurality of voices different in voice velocity by performing compression/ expansion processing while comparing the waveform of voice velocity on a display device.

SOLUTION: An input information control part 102 retrieves a voice file 103 by an inputted voice film name and delivers two voice file data to be connected to a display information control part 105. The display information control part 105 displays the velocity waveform of first half and second half voice file data respectively in the velocity waveform areas L, R of a display device 104, and gradually increases or decreases the velocity to adjust so that the velocity of the last part of the area L and that of the first part of the area R become equal. The set velocity increase/decrease multiplying factor information is delivered to a voice compression/expansion processing part 106, and two voice file data are connected as one voice file data.

Data supplied from the esp@cenet database - I2

(19)日本国特許庁 (JP)

(12)公開特許公報 (A)

(11)特許出願公開番号

特開平9-258778

(43)公開日 平成9年(1997)10月3日

(51) Int. Cl.⁶
G 10 L 3/02

識別記号

府内整理番号

F I

G 10 L 3/02

技術表示箇所
A

審査請求 未請求 請求項の数 3 OL (全 9 頁)

(21)出願番号 特願平8-69560
 (22)出願日 平成8年(1996)3月26日

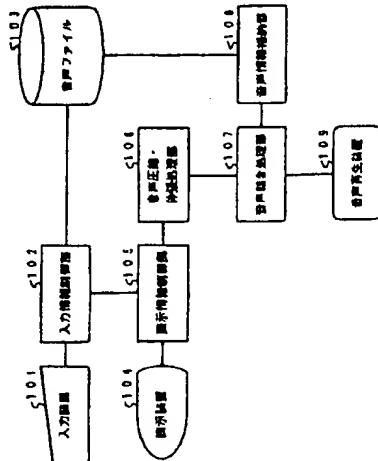
(71)出願人 000152985
 株式会社日立情報システムズ
 東京都渋谷区道玄坂1丁目16番5号
 (72)発明者 重岡 朋憲
 東京都渋谷区道玄坂一丁目16番5号 株式
 会社日立情報システムズ内
 (72)発明者 松井 功
 東京都渋谷区道玄坂一丁目16番5号 株式
 会社日立情報システムズ内
 (74)代理人 弁理士 磯村 雅俊 (外1名)

(54)【発明の名称】音声処理装置

(57)【要約】

【課題】 音声速度の異なる複数のサブナレーションから、スムーズに結合された一つのナレーションを生成することを可能とした音声処理装置を提供すること。

【解決手段】 音声波形のピーク値の平均間隔時間によって定義される音声速度の波形(速度波形)を表示装置に表示する表示手段104と、入力装置101から指示された情報(速度波形の変化度合いを示す加減速倍率を指定する情報と速度波形を変化させる時間を指定する情報)に基づいて前記表示された速度波形を変化させる速度波形変化手段101, 105と、前記速度波形変化手段による速度波形の変化に伴い前記音声情報を圧縮・伸張処理する圧縮・伸張手段106と、圧縮・伸張処理された前記複数の音声情報を結合する音声結合手段107からなる。



【特許請求の範囲】

【請求項1】 音声速度が相異なる複数の音声情報を結合する音声処理装置であって、音声波形のピーク値の平均間隔時間によって定義される音声速度の波形（速度波形という）を表示装置に表示する表示手段と、入力装置から指示された情報に基づいて前記表示された速度波形を変化させる速度波形変化手段と、前記速度波形変化手段による速度波形の変化に伴い前記音声情報を圧縮・伸張処理する圧縮・伸張手段と、圧縮・伸張処理された前記複数の音声情報を結合する音声結合手段を有することを特徴とする音声処理装置。

【請求項2】 請求項1記載の音声処理装置において、前記入力装置から指示された情報は、速度波形の変化度合いを示す加減速倍率を指定する情報と、速度波形を変化させる時間を指定する情報であることを特徴とする音声処理装置。

【請求項3】 請求項1または2記載の音声処理装置において、前記速度波形変化手段は、結合する複数の音声情報の速度を示す速度波形を前記表示装置にならべて表示し、前記入力装置から指定された情報によって該表示された複数の速度波形の結合部分の位置を合わせることにより、前記複数の音声情報における結合部分の音声速度を互いに近づける手段であることを特徴とする音声処理装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、音声速度の異なる複数のナレーションを1つのナレーションとして結合する際、複数のナレーションそれぞれの音声速度を表示装置で比較しながら、音声速度に関する適切な情報を入力し、圧縮・伸張処理を行うことにより、ナレーション間をスムーズに結合する音声処理装置に関する。

【0002】

【従来の技術】 従来、例えば一連の文章を朗読してこれを音声情報として記録する場合、この文章の量が非常に長い時には、複数回に分けて朗読した音声情報（以下、サブナレーションという）を、最終的に結合して一つのナレーションを作成する処理が行われる。このように、複数のサブナレーションを一つのナレーションとして結合する際、各サブナレーションは、朗読した日時等、環境が変化しているため、当然、同一人物が朗読したとしても、その音声速度（朗読速度、発声速度）が変化している。なお、本明細書における音声速度とは、単位時間あたりの、音声波形の振幅のピーク間の平均時間に対応した値を指している。

【0003】 このため、ナレーションを作成するにあたっては各サブナレーションを単に結合しただけでは、全体としてのつながり、バランスが不自然となってしまう。このため、各サブナレーションごとに音声速度のバランスをそろえた後、それを一つに結合する処理を

行う必要がある。ここで特に問題となるのが、各サブナレーション間の結合部分の音声速度を変更してそろえる処理である。従来、この場合の音声速度の変更は、オペレータが、実際に音声を聞いた感覚に基づいて音声を圧縮・伸張する部分を判断することによって行われている。

【0004】 音声速度を制御する音声処理方法としては、処理対象の音声と希望する出力音声との時間長の比によって、圧縮・伸張処理を行う方法がある。このうちの一技術として、「音声発声速度制御方法」（NTTデータ通信（株）、特開平05-241598号公報）が知られている。この従来技術は、音声を希望する時間長に変更する際、音声を1ピッチ単位に切り出して圧縮・伸張処理を行い、スムーズに各音声をつなぐことにより、品質の低下を防ぐようにした音声処理方法である。

【0005】

【発明が解決しようとする課題】 上述したように、複数のサブナレーションを結合して一つのナレーションを作成するにあたって、特に問題となるのが各サブナレーション間の結合部分の音声速度がなるべく均一にスムーズに繋がるようにするための処理である。これに対して、上記従来技術は、本発明でいうところの一つのナレーションの中で1ピッチ単位に音声を切り出してこの切り出した音声に対して圧縮・伸張処理を行う技術なので、上述のように朗読する環境の変化により大幅に音声速度が異なる可能性のある複数のサブナレーションを結合する場合、結合部周辺の音声が不自然にならないようにするためにには音声速度をどのように変更したらよいかなどについて全く配慮されておらず、これを行うためには、相変わらず人間の感覚に頼らなければならないという問題がある。本発明の目的は、従来の技術における上述の如き問題を解消し、音声速度の異なる複数音声からスムーズかつバランスよく結合された1つの音声を生成すること、さらに具体的には、複数のサブナレーションを1つの自然なナレーションに結合することを可能とした音声処理装置を提供することにある。

【0006】

【課題を解決するための手段】 本発明は、上記目的を達成するために、音声速度が相異なる複数の音声情報を結合する音声処理装置であって、音声波形のピーク値の平均間隔時間によって定義される音声速度の波形（速度波形という）を表示装置（104）に表示する表示手段と、入力装置（101）から指示された情報に基づいて前記表示された速度波形を変化させる速度波形変化手段（表示情報制御部105）と、速度波形変化手段による速度波形の変化に伴い前記音声情報を圧縮・伸張処理する圧縮・伸張手段（音声圧縮・伸張処理部106）と、圧縮・伸張処理された前記複数の音声情報を結合する音声結合手段（音声結合処理部107））を有していることを特徴としている。また、前記入力装置（101）か

ら指示された情報は、速度波形の変化度合いを示す加減速倍率を指定する情報（図3の305, 306）と、速度波形を変化させる時間を指定する情報（図3の307, 308）であることを特徴としている。また、前記速度波形変化手段は、結合する複数の音声情報の速度を示す速度波形を前記表示装置（104）にならべて表示し、入力装置（101）から指定された情報によって該表示された複数の速度波形の結合部分の位置を合わせることにより、複数の音声情報における結合部分の音声速度を互いに近づける手段であることを特徴としている。

【0007】

【発明の実施の形態】本発明では、音声波形のピーク値の平均間隔時間によって定義される音声速度の波形（速度波形という）を表示装置に表示し、入力装置から指示された情報（速度波形の変化度合いを示す加減速倍率を指定する情報と速度波形を変化させる時間を指定する情報）に基づいて表示された速度波形を変化させ、それに伴い音声情報を圧縮・伸張処理する圧縮・伸張して1つの音声情報に結合するようにしたものである。本発明によると、速度の制御を画面で視覚的に、かつ対話処理によって行うことが可能になり、必要な部分のみ変更を行うことができるとともに、作業時間の短縮や、音声品質の維持が可能となる。また、音声速度の変更を行った後、必要に応じて実際に結合された音声を聞くこともできるので、ナレータの経験や体調、環境の変化などによる速度のバラツキがあるサブナレーションを一つの自然なナレーションとして生成することが可能になる。

【0008】以下、本発明の実施例を図面を用いて詳細に説明する。図1は、本発明の一実施例に係る音声処理装置の構成を示すブロック図である。同図において、101はキーボードなどの入力装置、102は入力情報制御部、103は音声ファイル、104はCRTなどの表示装置、105は表示情報制御部、106は音声圧縮・伸張処理部、107は音声結合処理部、108は音声情報格納部、109は音声再生装置である。

【0009】図2は、本発明の一実施例に係る表示情報制御部105が行う処理フローを示す図である。図3は、本発明の一実施例の表示装置104に表示される速度情報定義画面を示す図である。同図において、301は速度波形領域L、302は速度波形領域R、303, 304は結合ファイル名欄、305, 306は加減速処理の倍率欄、307, 308は加減速処理の時間欄、309, 310は等速処理範囲の開始時間欄、311, 312は等速処理範囲の終了時間欄、313, 314は等速処理の倍率欄、315はプレビューボタン、316は確認ボタンを示している。図4は、本発明の一実施例に係る音声のサンプリングデータを示す図である。図5は、本発明の一実施例に係る速度波形領域を示す図である。図6は、本発明の一実施例に係る加減速処理による速度波形領域Lの波形移動を示す図である。図7は、本

発明の一実施例に係る加減速処理による速度波形領域Rの波形移動を示す図である。図8は、本発明の一実施例に係る等速処理による速度波形領域Rの波形移動を示す図である。

【0010】次に、図1、図3、および図4を用いて本実施例の音声処理装置の具体的な動作を詳細に説明する。まず、図3に示した表示装置104における速度情報定義画面の「結合ファイル名欄303, 304」に、入力装置101（キーボードなど）を用いて、結合する2つの音声ファイル名を入力する。入力情報制御部102は、入力された音声ファイル名によって音声ファイル103を検索し、検索された結合すべき2つの音声ファイルデータを表示情報制御部105にわたす。以降、結合する2つの音声ファイルデータのうち前半の音声ファイルデータを第1の音声ファイルデータ、後半の音声ファイルデータを第2の音声ファイルデータということにする。

【0011】表示情報制御部105では、入力情報制御部102からわたされた第1の音声ファイルデータおよび第2の音声ファイルデータの速度波形を、それぞれ表示装置104における速度情報定義画面の速度波形領域L301および速度波形領域R302に表示する。表示する領域の長さは、入力装置102によって、第1の音声ファイルデータについては加減速処理の時間欄307に、第2の音声ファイルデータについては加減速処理の時間欄308に指定入力する。

【0012】速度波形領域L301および速度波形領域R302の具体的な例を図5に示す。図5において、速度波形領域L301および速度波形領域R302のそれぞれに縦軸を平均時間、横軸を時間として速度波形が表示される。速度波形領域Lには結合する第1の音声ファイルデータすなわち前半のサブナレーションの最後の部分が、速度波形領域Rには結合する第2の音声ファイルデータすなわち後半のサブナレーションの最初の部分が表示される。

【0013】ここで、図5の速度波形領域L301および速度波形領域R302の縦軸の値となる平均時間の算出方法を図4を用いて説明する。図4に示したように、サンプリングデータ中の音声波形の振幅が大きいところをピーク位置として検出し、ピーク間の時間を計測する。そして、速度波形領域に表示するために、単位時間あたりの平均時間を算出し、その結果、速度波形領域に速度が表示される。なお、本明細書における音声速度とは、単位時間あたりのピーク間の平均時間を指している。

【0014】すなわち、図5においては、速度波形領域Lのグラフが右下がりになっているので、朗読している速度が徐々に遅くなっていることを示し、逆に、速度波形領域Rのグラフは、速度波形領域Rは、速度波形領域Lに示された最後（平均時間が0の位置での点）の音声

速度よりも朗読時間が最初速く始まり、朗読速度は更に徐々に速くなっていることを示している。

【0015】次に、速度波形領域Lの最後の部分の速度および速度波形領域Rの最初の部分の速度を比較し、速度が異なっている場合に加減速処理を行う。加減速処理とは第1の音声ファイルデータの最後の部分および第2の音声ファイルデータの最初部分、つまり結合部分の速度を調整してスムーズに結合するもので、徐々に速度を上げたりあるいは徐々に速度を下げたりする処理を行う。

【0016】速度を上げる場合には、入力装置101により加減速処理の倍率欄305または306に1より小さな値を入力するか、あるいは、加減速処理の倍率欄305または306の脇に設けられた上矢印をマウスでクリックすることにより行う。また、速度を下げる場合には、入力装置101により加減速処理の倍率欄305または306に1より大きな値を入力するか、あるいは、加減速処理の倍率欄305または306の脇に設けられた下矢印をマウスでクリックすることにより行う。これらの操作により速度波形が移動し、速度波形領域Lの最後の部分および速度波形領域Rの最初の部分の速度が等しくなるように調整される。ここで、処理の確認のため、プレビューボタン315をクリックすると、設定した加減速倍率情報が表示情報制御部105から音声圧縮・伸張処理部106にわたされ、2つの音声ファイルデータを1つの音声ファイルデータとして結合される。そして、結合された音声ファイルデータは必要に応じて音声再生装置109により再生される。

【0017】以上の処理が終了した後、今度は音声ファイル全体を速度波形領域Lおよび速度波形領域Rに並べて表示する。この波形を見て、特に速い部分と遅い部分がある場合に次に説明する等速処理を行う。

【0018】等速処理とは、指定範囲の速度を同じ比率で変化させる処理である。指定範囲とは、入力装置101により等速処理範囲の開始時間欄309、310および終了時間欄311、312に値を入力するか、あるいは、マウスにより速度波形領域をドラッグして指定する。速度を上げる場合には、入力速度101により等速処理の倍率欄313、314に1より小さな値を入力するか、あるいは、等速処理の倍率欄313、314の脇に設けられた上矢印をマウスでクリックすることにより行う。また、速度を下げる場合には、入力装置101により等速処理の倍率欄313、314に1より大きな値を入力するか、あるいは、等速処理の倍率欄313、314の脇に設けられた下矢印をマウスでクリックすることにより行う。これらの操作により速度波形が移動し、指定範囲の速度を変化させ、全体の速度バランスをそろえる。

【0019】ここで、再度、音声を確認するため、プレビューボタン315をクリックすると、設定した等速情

報が表示情報制御部105から音声圧縮・伸張処理部106にわたされ、音声の圧縮・伸張を行う。処理後の音声ファイルデータが音声結合処理部107にわたされ、2つの音声ファイルデータを1つの音声ファイルデータとして結合する。そして、結合された音声ファイルデータを音声再生装置109により再生する。その結果、不都合な場合は上記加減速処理、または、等速処理を繰り返し、適切な速度調整を行う。最後に、速度調整の確認が終了したら、確認ボタン316をクリックすることにより、音声結合処理部107から音声情報格納部108に結合された音声ファイルデータがわたされ、音声ファイル103に格納される。

【0020】次に、図2に示したフローチャートに基づいて、上記表示情報制御部105および音声圧縮・伸張処理部106の動作を詳細に説明する。まず、表示情報制御部105において、入力装置101より入力された結合すべき2つの音声ファイル名欄303の内容を読み込む(ステップ201)。次に、それぞれの音声ファイルデータ中のピーク位置を検出し(ステップ202)、ピーク間の時間長を計測する(ステップ203)。計測されたピーク間の時間長を単位時間(ここでは1秒)で区切り、平均時間を算出する(ステップ204)。ステップ204をもう少し分かり易くいうと、例えば、速度波形領域Lに表示を行うとき、横軸を1秒単位にすると仮定すると、図4に示した「ピーク間の時間」が、その1秒の間に何度も現れるが、平均時間の算出とはそれら「ピーク間の時間」の平均を求めるに相当する。

【0021】そして、単位時間あたりの上記平均時間をポイントし、ポイント間をスムーズにつなげて速度波形領域L301に速度波形を表示する(ステップ205、図5参照)。なお、速度波形領域R302も同様の処理を行う。

【0022】次に、第1の音声ファイルデータと第2の音声ファイルデータの結合部分に速度調整を行う必要がある場合には加減速処理を行う(ステップ206:YES、ステップ207～ステップ211のループ)。次に、加減速処理の一実施例を図6および図7を用いて説明する。図6は第1の音声ファイルデータを表している速度波形領域Lの最後の部分を示し、図7は第2の音声ファイルデータを表している速度波形領域Rの最初の部分を示している。速度波形領域Lおよび速度波形領域Rに表示される領域の時間長は加減速処理の時間欄307、308に入力する値により変更が可能である(ステップ208)。例えば、加減速処理の時間欄307に10を入力した場合、速度波形領域Lには、第1の音声ファイルデータの最後の部分10秒が表示され、加減速処理の時間欄308に10を入力した場合、速度波形領域Rには、第2の音声ファイルデータの最初の部分10秒が表示される。

【0023】前述のように加減速処理の時間欄307、

308に10を入力した場合にはそれぞれ表示される時間長は10秒間であり、図6の速度波形領域Lに表示された速度波形の前端の値をL2、後端の値をL1とし、図7の速度波形領域Rに表示された速度波形の前端の値をR1、後端の値をR2とする。図6の速度波形領域Lの最後の部分L1と図7の速度波形領域Rの最初の部分R1の速度がそろっていない場合、これをそろえるために加減速処理を行う。図6および図7に示した実施例では、速度波形領域Lの最後の部分L1を120ミリ秒、L2を80ミリ秒とし、R1を40ミリ秒、R2を15ミリ秒である。この場合、速度波形領域Lの最後の部分L1に対して加速処理、速度波形領域Rの最初の部分R1に対して減速処理を行う。

【0024】次に加減速処理倍率の入力（ステップ209）を行う。速度波形領域Lの最後の部分L1に対して加速処理を行うには、入力装置101により加減速処理の倍率欄305に1より小さな値を入力するか、あるいは、加減速処理の倍率欄305の脇に設けられた上矢印をマウスでクリックすることにより値を少しづつ（例えば0.01ずつに設定）減少させることによって行う。速度波形領域Rの最初の部分R1に対して減速処理を行うには、入力装置101により加減速処理の倍率欄308に1より大きな値を入力するか、あるいは、加減速処理の倍率欄308の脇に設けられた下矢印をマウスでクリックすることにより値を少しづつ（例えば0.01ずつに設定）増加させることによって行う。

【0025】上述したステップ208における加減速処理時間の入力およびステップ209における加減速処理倍率の入力によって図6および図7の速度波形の移動が実行される（ステップ210）。次に、速度波形が移動する仕組みについて説明する。速度波形領域LのL1～L2間を所定時間（例えば2秒）あたりに等間隔に分割し、ファイルの最後の部分L1が含まれる領域を分割数の最大値を番号（ここでは5）とし、L2の方に1ずつ減少させながら分割番号を順次割り当てる。よって、L2の分割番号は1となる。速度波形領域Rも同様に、ファイルの最初の部分R1が含まれる領域を分割数の最大値を番号（ここでは5）とし、R2の方に1ずつ減少させながら分割番号を順次割り当てる。よって、R1の分割番号は5、R2の分割番号は1となる。

【0026】次に、分割番号ごとの領域における平均時間の算出方法を説明する。

分割番号領域平均時間=分割番号領域処理倍率×分割番号領域平均時間

分割番号領域処理倍率=1+(加減速処理倍率-1)×分割番号/分割数

【0027】従って、例えば、速度波形領域Lの加減速処理の倍率を0.5とした場合の分割番号ごとの領域の平均時間は次のようになる。

$$\text{新L1} = (1 + (0.5 - 1) \times 5 / 5) \times 120 = 6$$

$$\text{新L2} = (1 + (0.5 - 1) \times 1 / 5) \times 80 = 72$$

【0028】また、速度波形領域Rの加減速処理の倍率を1.5とした場合の分割番号ごとの領域の平均時間は次のようになる。

$$\text{新R1} = (1 + (1.5 - 1) \times 5 / 5) \times 40 = 60$$

$$\text{新R2} = (1 + (1.5 - 1) \times 1 / 5) \times 15 = 16.5$$

【0029】これらの処理の結果、図6に示すように、速度波形領域Lの速度波形（L1～L2）が上方に移動し（新L1～新L2）、速度が徐々に速くなる、つまり、平均時間の短縮が行われる。また、図7に示すように、速度波形領域Rの速度波形（R1～R2）が下方に移動し（新R1～新R2）、速度が徐々に遅くなる、つまり、平均時間の拡張が行われる。平均時間の短縮とは、例えば、図4におけるピーク間の時間を狭くすることであり、拡張とは逆にピーク間の時間を広くすることである。ここで、プレビューボタン315をクリックすると、音声の再生が行われ、問題がなければ確認ボタン316をクリックする。問題がある場合には上記ステップ207～ステップ211の処理を問題がなくなるまで繰り返す。

【0030】次に、音声速度のバランスが合わない部分が存在する場合、等速処理を行う（ステップ212；YES、ステップ213～ステップ217のループ）。等速処理は次のようにして行われる。まず、等速処理範囲の開始時間欄309、310（ここでは20秒とする）および終了時間欄311、312（ここでは30秒とする）に入力する（ステップ214）。さらに、加減速処理と同様に、等速処理倍率欄313、314（ここでは1.5倍とする）に入力する（ステップ215）。その結果、図8に示すように、速度波形領域Rが下方に移動し、速度が遅くなる、つまり、平均時間の拡張が行われる（ステップ216）。ここで、プレビューボタン309をクリックすると、音声の再生が行われ、問題がなければ確認ボタン310をクリックする。問題がある場合にはステップ213～ステップ217のループの繰り返し処理を行うか、あるいは再度、ステップ206に戻って加減速処理を行う。上記加減速処理および等速処理による音声速度の変更を、2つの音声速度がスムーズにかつバランスよくつながるまで繰り返す。音声ファイルデータが他にも存在する場合には、上記加減速処理および等速処理を必要なだけ繰り返す。

【0031】以上説明した如く、本実施例によれば、音声速度の異なる複数の音声ファイルデータをスムーズでバランスのよい均一な音声速度で結合することが可能に

なる。なお、上記実施例は、本発明の一例を示したものにすぎず、加減速倍率、時間などのパラメータや表示画面例などは、本例に限定されるべきものではなく、各種の変更が可能であることはいうまでもない。

[0032]

【発明の効果】以上、詳細に述べた如く、本発明によれば、音声速度の変更制御を対話処理によって行うことにより、複数の音声ファイルデータをスムーズに合成するための作業時間を短縮することが可能となる。また、ナレータの経験や体調、環境の変化などによる、速度のバラツキがある複数のサブナレーションを一つの自然なナレーションとして生成することが可能となる。

【図面の簡単な説明】

【図1】本発明の一実施例に係る音声処理装置の構成を示すブロック図である。

【図2】本発明の一実施例に係る表示情報制御部105が行う処理フローを示す図である。

【図3】本発明の一実施例に係る速度情報定義画面を示す図である。

【図4】本発明の一実施例に係る音声のサンプリングデータ（振幅／時間）を示す図である。

【図5】本発明の一実施例に係る速度波形領域を示す図である。

【図6】本発明の一実施例に係る加減速処理による速度波形領域Lの波形移動を示す図である。

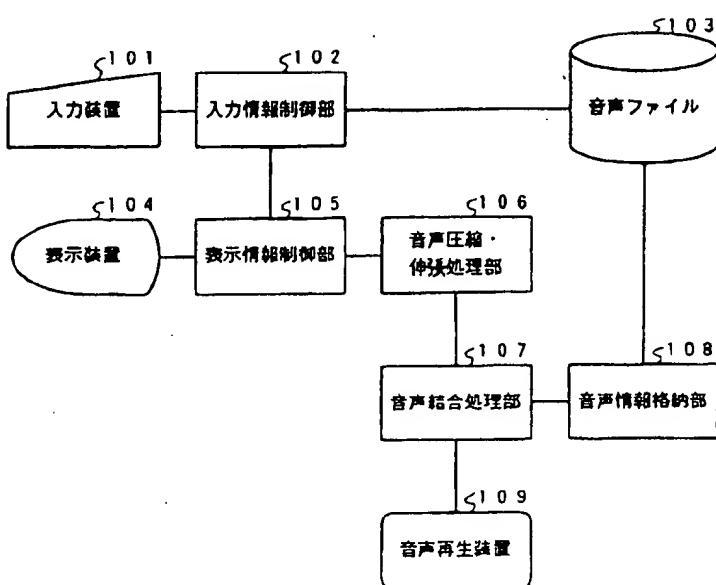
【図7】本発明の一実施例に係る加減速処理による速度波形領域Rの波形移動を示す図である。

【図8】本発明の一実施例に係る等速処理による速度波形領域Rの波形移動を示す図である。

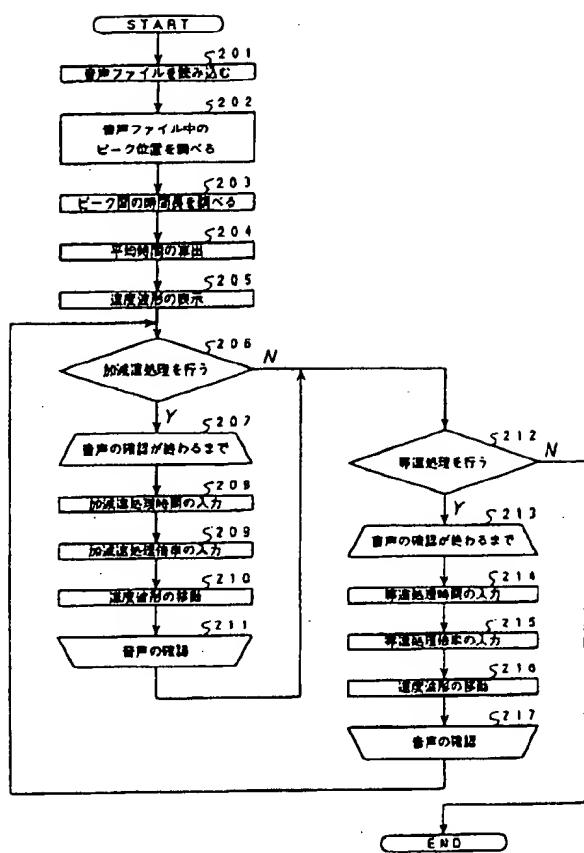
【符号の説明】

101：入力装置、102：入力情報制御部、103：音声ファイル、104：表示装置、105：表示情報制御部、106：音声圧縮・伸張処理部、107：音声結合処理部、108：音声情報格納部、109：音声再生装置、301：速度波形領域L、302：速度波形領域R、303、304：結合ファイル名欄、305、306：加減速処理の倍率欄、307、308：加減速処理の時間欄、309、310：等速処理範囲の開始時間欄、311、312：等速処理範囲の終了時間欄、313、314：等速処理の倍率欄、315：プレビューボタン、316：確認ボタン

【図1】



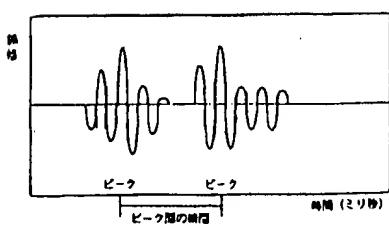
【図2】



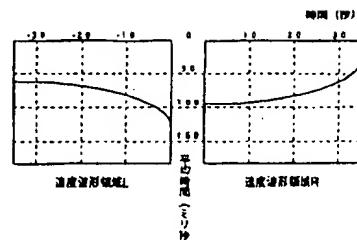
【図3】

<p>速度波形領域L</p> <p>ファイル名 <input type="text"/> S303</p> <p>加速減速処理</p> <p>倍率 <input type="text"/> S305</p> <p>時間 <input type="text"/> S307</p> <p>等速処理 <input type="text"/> 309 <input type="text"/> 311</p> <p>範囲 <input type="button"/> 開始 <input type="button"/> 終了 <input type="button"/></p> <p>倍率 <input type="text"/> S313</p>	<p>速度波形領域R</p> <p>ファイル名 <input type="text"/> S304</p> <p>加速減速処理</p> <p>倍率 <input type="text"/> S306</p> <p>時間 <input type="text"/> S308</p> <p>等速処理 <input type="text"/> 310 <input type="text"/> 312</p> <p>範囲 <input type="button"/> 開始 <input type="button"/> 終了 <input type="button"/></p> <p>倍率 <input type="text"/> S314</p>
<input type="button"/> 315 プレビュー <input type="button"/> 316	

【図4】



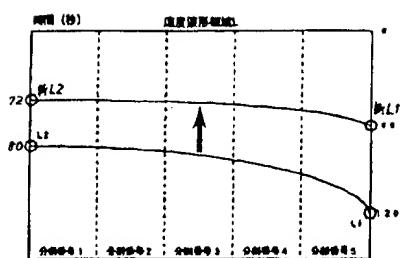
【図5】



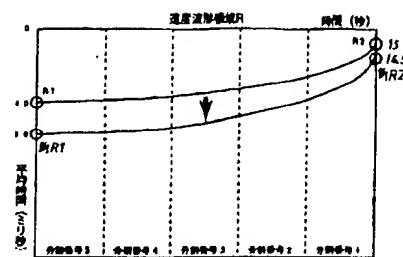
(9)

特開平9-258778

【図6】



【図7】



【図8】

